

การศึกษาเปรียบเทียบประสิทธิภาพการจำแนกภาวะกระดูกหักจากภาพเอกซเรย์ กระดูกข้อมือด้วยโครงข่ายประสาทคอนโวลูชัน

A Comparative Study of Performance in Bone Fracture Classification from
Extremity X-ray Images using Convolutional Neural Networks (CNN)

พงษ์ศธร เชิดสม

Pongsathorn Chedsom

คณะสถาปัตยกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

Faculty of Architecture, Khon Kaen University

E-mail: pongche@kku.ac.th

(Received: 13 August 2025, Revised: 29 May 2026, Accepted: 6 June 2026)

<https://doi.org/10.57260/stc.2026.1100>

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบประสิทธิภาพการจำแนกภาวะกระดูกหักจากภาพเอกซเรย์กระดูกข้อมือโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันจากชุดข้อมูล Bone Fracture Multi-Region X-ray จำนวนทั้งหมด 10,432 ภาพ แบ่งเป็นภาพกระดูกหักและไม่หักอย่างละจำนวน 5,216 ภาพ แบ่งสัดส่วนข้อมูลสำหรับการเรียนรู้ การตรวจสอบ และการทดสอบแบบจำลองเป็นร้อยละ 70, 20 และ 10 ตามลำดับ กระบวนการทดลองแบ่งเป็น 2 ระยะ ได้แก่การประเมินแบบจำลองพื้นฐาน 6 ชนิดเพื่อสกัดคุณลักษณะเด่น และการปรับปรุงประสิทธิภาพขั้นเชื่อมโยงแบบสมบูรณ์ รวม 60 รูปแบบ ผลการศึกษาพบว่าแบบจำลอง VGG16 และ MobileNet ให้ประสิทธิภาพสูงสุดในกลุ่มแบบจำลองพื้นฐาน มีค่าความถูกต้องร้อยละ 99.450 ค่าความสูญเสีย 0.030 และเมื่อนำมารวมกับขั้นเชื่อมโยงแบบสมบูรณ์ในรูปแบบที่เหมาะสม (Group 4 + Pattern 3) สามารถเพิ่มค่าความถูกต้องเป็นร้อยละ 100 และลดค่าความสูญเสียลงได้อย่างมีนัยสำคัญเมื่อเปรียบเทียบกับแบบจำลองพื้นฐาน โดยค่าความถูกต้องนี้เป็นผลลัพธ์เฉพาะชุดข้อมูลที่ใช้ในการทดสอบครั้งนี้ งานวิจัยนี้แสดงให้เห็นถึงศักยภาพในการพัฒนาเป็นเครื่องมือช่วยรังสีแพทย์เพื่อสนับสนุนการวินิจฉัยเบื้องต้น ซึ่งจะช่วยลดภาระงานและเพิ่มความรวดเร็วในการคัดกรองเบื้องต้นได้อย่างมีประสิทธิภาพ

คำสำคัญ: การจำแนกภาวะกระดูกหัก ภาพเอกซเรย์กระดูกข้อมือ โครงข่ายประสาทคอนโวลูชัน การจำแนก
ภาพถ่ายทางการแพทย์ การศึกษาเปรียบเทียบประสิทธิภาพ

Abstract

This research aims to compare the performance of bone fracture classification from extremity X-ray images using Convolutional Neural Networks (CNN). A total of 10,432 images from the Bone Fracture Multi-Region X-ray dataset were utilized, partitioned into training (70%), validation (20%), and test sets (10%). The methodology consisted of two phases: evaluating six baseline models for feature extraction and optimizing the Fully Connected Layer through 60 variations. The results indicated that the VGG16 and MobileNet model achieved the highest performance among the baselines. By integrating it with an optimized Fully Connected structure (Group 4 + Pattern 3), the model reached an accuracy of 100% with a significant reduction in loss compared to the baseline models. While this perfect accuracy is specific to the characteristics of the current dataset, the findings demonstrate the potential of the developed model as a diagnostic aid for radiologists. This research demonstrates the potential for development into a tool that assists radiologists in preliminary diagnosis, thereby effectively reducing workload and accelerating initial screening.

Keywords: Bone fracture classification, Extremity X-ray, Convolutional neural network, Medical image classification, Comparative performance analysis

บทนำ

การวินิจฉัยภาวะกระดูกหักจากภาพเอกซเรย์ (X-ray) โดยเฉพาะบริเวณกระดูกซี่โครงหักจากภาพเอกซเรย์กระดูกปลาย (Extremity X-ray images) เช่น แขน ขา ข้อมือ และข้อเท้า เป็นสิ่งสำคัญอย่างยิ่งในการดูแลผู้ป่วย เพราะกระดูกซี่โครงหักอาจบ่งชี้ถึงการบาดเจ็บรุนแรงต่ออวัยวะภายในช่องอก เช่น ปอดฉีกขาด เลือดออกในช่องอก หรือการบาดเจ็บต่อหัวใจและหลอดเลือดใหญ่ การวินิจฉัยที่รวดเร็วและแม่นยำจะช่วยให้แพทย์สามารถวางแผนการรักษาได้อย่างเหมาะสมและรวดเร็ว หากตรวจพบรอยแตกหักล่าช้าหรือผิดพลาดอาจนำไปสู่ภาวะแทรกซ้อนเรื้อรัง การผิดรูปของกระดูก หรือความพิการที่ส่งผลต่อการใช้ชีวิตของผู้ป่วย ซึ่งช่วยลดภาวะแทรกซ้อนและอัตราการเสียชีวิตของผู้ป่วยได้ การตรวจวินิจฉัยภาวะกระดูกหักจากภาพเอกซเรย์กระดูกปลายนั้นมีข้อจำกัดหลายประการ เช่น ความยากในการมองกระดูกซี่โครงมีลักษณะโค้งและซับซ้อน ทับซ้อนกับเงาของอวัยวะอื่น ๆ เช่น ปอด หัวใจ และกระบังลม ส่งผลต่อการมองเห็นรอยร้าวหรือการเคลื่อนที่ของกระดูกเป็นไปได้ยาก โดยเฉพาะอย่างยิ่งในกรณีของกระดูกหักที่ไม่สมบูรณ์หรือมีขนาดเล็ก ความเชี่ยวชาญในการวิเคราะห์และการวินิจฉัยกระดูกหักจากภาพเอกซเรย์ต้องอาศัยประสบการณ์และความเชี่ยวชาญของผู้ที่เกี่ยวข้อง รวมถึงจำนวนของภาพเอกซเรย์ที่ต้องอ่านและวินิจฉัยส่งผลต่อระยะเวลาในการวินิจฉัยและความแม่นยำในการวินิจฉัย

การวินิจฉัยความผิดปกติของกระดูก (Bone abnormalities) จากภาพเอกซเรย์ชุดข้อมูล Mura-v1.1 ของ (Navale. S., 2023) จำนวน 42,000 ภาพ แบ่งชุดข้อมูลออกเป็นร้อยละ 80 สำหรับการเรียนรู้ของแบบจำลองและร้อยละ 20 สำหรับการทดสอบแบบจำลองและใช้แบบจำลอง VGG16 ที่ผ่านการเรียนรู้ล่วงหน้า (Pre-training) ในการเรียนรู้ของแบบจำลองพบว่าให้ความถูกต้องร้อยละ 85.82

การวินิจฉัยภาวะกระดูกหักของ (Sahin, 2023) โดยการตรวจจับและการจำแนกโดยจากภาพเอกซเรย์ทรวงอกแบ่งภาพออกเป็นภาวะกระดูกหักจำนวน 71 ภาพ และไม่เป็นภาวะกระดูกหักจำนวน 105 ภาพ แบ่งชุดข้อมูลออกเป็นร้อยละ 70 สำหรับการเรียนรู้ของแบบจำลองและร้อยละ 30 สำหรับการทดสอบแบบจำลอง และใช้เทคนิคการแบ่งข้อมูล K-fold 10 ชุดข้อมูล จำนวน 12 รูปแบบ/เทคนิค ได้แก่ Gaussian NB, Multinomial NB, Extra trees classifier, LGBM classifier, Ada boost classifier, Gradient boosting classifier, MLP classifier, XGB classifier, Quadratic discriminant analysis, Random forest classifier, Logistic regression, Linear discriminant analysis และใช้เทคนิค Grid search ในการทดลองพบว่า Linear discriminant analysis ให้ความถูกต้องมากที่สุดร้อยละ 88.67 และ AUC 0.89

การตรวจจับและจำแนกกระดูกต้นแขนหักของ (Barua et al., 2023) โดยใช้ชุดข้อมูล MURA (Musculoskeletal radiographs) จำนวน 40,561 ภาพ ใช้แบบจำลองเดี่ยวในการทดลองได้แก่ CNN, VGG16, VGG19, DenseNet121, DenseNet169 พบว่ามีค่าความถูกต้องร้อยละ 80.00, 79.00, 80.00, 80.00, 82.00 ตามลำดับและแบบจำลองโดยใช้เทคนิคการรวมกลุ่ม (Ensemble) ได้แก่ Ensemble-1 (VGG16 และ VGG19), Ensemble-2 (DenseNet121 และ DenseNet169) พบว่ามีค่าความถูกต้องร้อยละ 83.00, 85.00 ตามลำดับ

การจำแนกกระดูกหักของข้อมือของ (Senapati et al., 2024) จากภาพเอกซเรย์จำนวน 200 ภาพ ขนาด 150x150 แบ่งภาพออกเป็นภาวะกระดูกหักจำนวน 120 ภาพ และไม่เป็นภาวะกระดูกหักจำนวน 80 ภาพ ด้วยแบบจำลอง 5 แบบจำลองได้แก่ VGG16, VGG19, InceptionV3, ResNet50, Proposed Model พบว่ามีค่าความถูกต้องร้อยละ 84.00, 78.00, 69.00, 84.00, 98.00 ตามลำดับ

การจำแนกและการตรวจจับกระดูกหักของ (Thota et al., 2024) จากภาพเอกซเรย์รังสีจำนวน 5,000 ภาพ ขนาด 256x256 แบ่งชุดข้อมูลออกเป็นร้อยละ 70 สำหรับการเรียนรู้ของแบบจำลองและร้อยละ 30 สำหรับการทดสอบแบบจำลอง แบ่งภาพออกเป็นภาวะกระดูกหักและไม่เป็นภาวะกระดูกหัก โดยใช้แบบจำลองที่พัฒนาเองจาก CNN และ MobileNet พบว่ามีค่าความถูกต้องร้อยละ 86.00 และ 98.00 ตามลำดับ

การจำแนกกระดูกหักจากภาพเอกซเรย์รังสีของกล้ามเนื้อและกระดูกของ (Islam et al., 2025) จากชุดข้อมูล MURA (Musculoskeletal radiographs) จำนวน 40,019 ภาพ แบ่งชุดข้อมูลออกเป็นร้อยละ 80 จำนวน 22,418 ภาพ สำหรับการเรียนรู้ของแบบจำลองและร้อยละ 20 จำนวน 5,600 ภาพ สำหรับการทดสอบแบบจำลอง โดยผ่านการปรับขนาด (Scaling) การครอบตัด (Cropping) และการเพิ่มข้อมูล (Augmentation) แบ่งออกเป็น 2 ประเภทได้แก่ "ปกติ" (ไม่มีกระดูกหัก) และ "ผิดปกติ" (มีกระดูกหัก) ด้วย

แบบจำลอง 4 แบบจำลอง ได้แก่ ViT-16, CNN, ResNet50 (Residual Networks), ResNet-34 พบว่ามีค่าความถูกต้องร้อยละ 88.00, 87.00, 86.00, 80.00 ตามลำดับ

ในปัจจุบันเทคโนโลยีปัญญาประดิษฐ์ (Artificial intelligence: AI) โดยเฉพาะโครงข่ายประสาทคอนโวลูชัน (Convolutional neural networks: CNN) ได้กลายเป็นเครื่องมือหลักที่มีประสิทธิภาพโดดเด่นในการประมวลผลภาพทางการแพทย์ เนื่องจากมีความสามารถในการเรียนรู้คุณลักษณะที่ซับซ้อนจากภาพได้โดยตรง และจากการทบทวนวรรณกรรมพบว่าการประยุกต์ใช้แบบจำลองที่หลากหลายรูปแบบเพื่อจำแนกภาวะผิดปกติของกระดูก จากการศึกษาวิจัยในช่วงปี 2022 - 2025 พบว่า การจำแนกความผิดปกติและภาวะกระดูกหักมีการพัฒนาอย่างต่อเนื่องโดยในกลุ่มงานที่ใช้ชุดข้อมูลขนาดใหญ่และโมเดลพื้นฐาน เช่น งานของ (Navale, 2023) และ (Barua et al., 2023) ที่ใช้ชุดข้อมูล MURA พบค่าความถูกต้องอยู่ในช่วงร้อยละ 79.00 - 85.82 ขณะที่งานวิจัยล่าสุดของ (Islam et al., 2025) ที่ใช้เทคนิค Visual transformers (ViT) สามารถเพิ่มระดับความถูกต้องขึ้นที่ร้อยละ 88.00 สำหรับกลุ่มงานวิจัยที่มุ่งเน้นพื้นที่เฉพาะส่วนหรือใช้แบบจำลองปรับปรุงใหม่ เช่น การจำแนกกระดูกข้อมือของ (Senapati et al., 2024) และการใช้ MobileNet ของ (Thota et al., 2024) แสดงให้เห็นถึงประสิทธิภาพที่เพิ่มมากขึ้น โดยสามารถให้ค่าความถูกต้องได้สูงสุดถึงร้อยละ 98.00 ซึ่งงานวิจัยที่ผ่านมาจะแสดงให้เห็นถึงประสิทธิภาพของอัลกอริทึมและแบบจำลองหลากหลายรูปแบบทั้งการใช้แบบจำลองมาตรฐาน (Baseline models) เทคนิคการรวมกลุ่ม (Ensemble) หรือเทคนิค Transformers เพื่อเพิ่มความถูกต้องเป็นหลัก จากการทบทวนวรรณกรรมพบข้อจำกัดสำคัญคือการขาดการออกแบบเชิงลึกเกี่ยวกับการปรับแต่งโครงสร้างชั้นเชื่อมต่อแบบสมบูรณ์ (Fully connected layer) เพื่อให้มีความเหมาะสมกับลักษณะเฉพาะของภาพเอกซเรย์กระดูกปลาย (Extremity x-ray) งานวิจัยนี้จึงมุ่งเน้นการศึกษาและเปรียบเทียบประสิทธิภาพโดยการปรับแต่งโครงสร้างชั้นเชื่อมต่อแบบสมบูรณ์ 60 รูปแบบ โดยใช้รูปแบบการปรับแต่งจำนวนโนดของชั้นเชื่อมต่อแบบสมบูรณ์จาก (Chedsom & Kanarkard, 2023) เพื่อหาโครงสร้างที่เหมาะสมที่สุดในการจำแนก (Classification) สำหรับชุดข้อมูล Bone fracture multi-region x-ray ร่วมกับการใช้งานแบบจำลองแบบพื้นฐานที่ใช้ในการสกัดคุณลักษณะเด่นของข้อมูล (Feature extraction) ซึ่งจะช่วยให้ประสิทธิภาพการจำแนกภาวะกระดูกหักให้สูงกว่ามาตรฐานของแบบจำลองพื้นฐานทั่วไป อีกทั้งยังแสดงให้เห็นถึงศักยภาพในการพัฒนาเป็นเครื่องมือสนับสนุนการวินิจฉัยเบื้องต้น (Preliminary diagnosis) เพื่อช่วยรังสีแพทย์ในการลดภาระงานและเพิ่มความรวดเร็วในการคัดกรองผู้ป่วยในสถานพยาบาลได้อย่างแม่นยำและน่าเชื่อถือในทางคลินิก

วัตถุประสงค์การวิจัย

เพื่อศึกษาและเปรียบเทียบประสิทธิภาพการจำแนกภาวะกระดูกหักจากภาพเอกซเรย์กระดูกปลายระหว่างแบบจำลองพื้นฐาน (Baseline models) 6 สถาปัตยกรรม ร่วมกับการปรับแต่งโครงสร้างชั้นเชื่อมต่อแบบสมบูรณ์ (Fully connected layer) 60 รูปแบบ

ระเบียบวิธีวิจัย

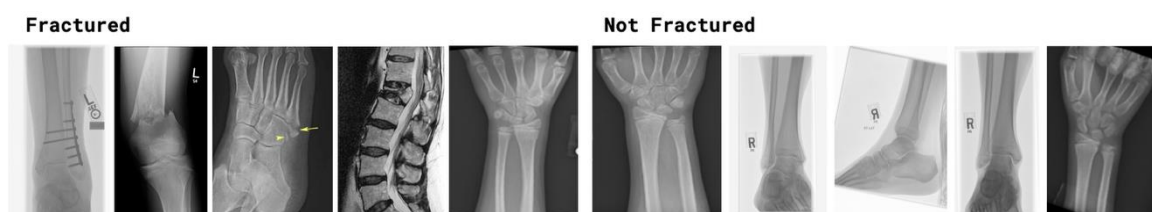
เครื่องมือและซอฟต์แวร์ที่ใช้ในการทดลอง

ในการทดลองนี้ได้ใช้เครื่องมือและซอฟต์แวร์ในการทดลองและพัฒนาแบบจำลองได้แก่

- GPU: Nvidia RTX 4070 SUPER (12GB)
- CPU: Intel Core i9-13900KS
- RAM: 32GB
- OS: Ubuntu 22.04 LTS
- Deep learning framework: keras (Tensor flow)
- Programming language: Python

ชุดข้อมูลในการทดลอง

ชุดข้อมูลที่นำมาใช้ในการทดลองจากเว็บไซต์ kaggle.com ของ (Rodrigo, 2025) เป็นชุดข้อมูลที่ได้รับคามนิยมในบทความต่าง ๆ มีผู้ดาวน์โหลดนำไปใช้งานโดยเฉลี่ยเดือนละ 8,856 ครั้ง ชุดข้อมูลประกอบไปด้วยภาพเอกซเรย์กระดูกซี่โครงที่หักและไม่หักครอบคลุมทุกส่วนของร่างกายทางกายวิภาค รวมถึงแขนขา ขาส้นบน เอว สะโพก เข่า เป็นต้น (Bone fracture multi-region) ทั้งหมด 10,581 ภาพ แบ่งเป็นกระดูกหัก (Fractured) จำนวน 5,181 ภาพ และไม่หัก (Not fractured) จำนวน 5,400 ภาพ ดังตัวอย่างแสดงในภาพที่ 1



ภาพที่ 1 ภาพจากชุดข้อมูล (Rodrigo, M., 2025)

การเตรียมชุดข้อมูล

จากชุดข้อมูลต้นฉบับได้เลือกสุ่มภาพประเภทละ 5,181 ภาพ จากจำนวนภาพที่น้อยที่สุดในแต่ละประเภทเพื่อให้ชุดข้อมูลทั้งสองประเภทมีความสมดุล (Balanced dataset) รวมจำนวน 10,362 ภาพ การจัดเตรียมชุดข้อมูลสำหรับการทดลองได้แบ่งชุดข้อมูลออกเป็นชุดข้อมูลสำหรับการสร้างการเรียนรู้ของแบบจำลอง Training set ร้อยละ 70 จำนวน 7,252 ภาพ ชุดข้อมูล Validation set ร้อยละ 20 จำนวน 2,072 ภาพ ชุดข้อมูล Test set ร้อยละ 10 จำนวน 1,038 ภาพ ดังแสดงในภาพที่ 2 (a) และใช้เทคนิคการแบ่งข้อมูลด้วย K-Fold cross validation (K=5) มาใช้ในการทดสอบประสิทธิภาพของแบบจำลองและเพื่อให้ข้อมูลมีการกระจายตัวอย่างสม่ำเสมอ ช่วยให้การทดสอบประสิทธิภาพของแบบจำลอง เช่น ค่าความถูกต้องหรือค่าความผิดพลาดของแบบจำลองความสอดคล้องกับข้อมูลทั้งหมดโดยใช้การจัดกลุ่มข้อมูลจากการสุ่มให้

เป็นชุดข้อมูลใหม่ทั้ง 5 กลุ่มดังแสดงในภาพที่ 2 (b) การแบ่งชุดข้อมูลสำหรับการทดสอบประสิทธิภาพของแบบจำลอง

การเตรียมข้อมูลเพื่อใช้ในการเรียนรู้ของแบบจำลองหลังจากการเตรียมชุดข้อมูลแล้วดังนี้

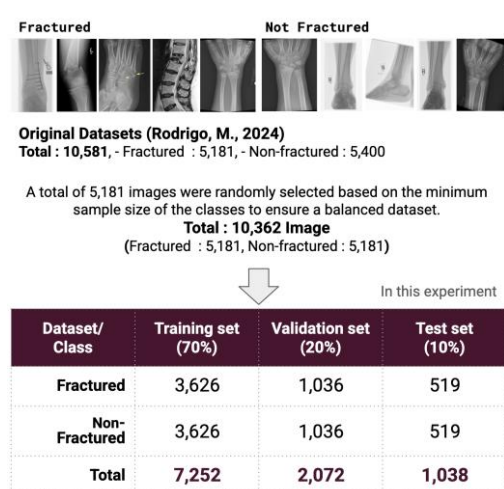
- ใช้ฟังก์ชัน Image data generator และ Flow_from_directory ของ Keras เพื่อจัดเตรียมโครงสร้างของโพลเดอร์ชุดข้อมูล แบ่งออกเป็นโพลเดอร์หลักได้แก่ Fractured สำหรับภาพกระดูกหัก และ Not fractured สำหรับภาพกระดูกไม่หัก ภายในโพลเดอร์หลักประกอบไปด้วยโพลเดอร์ย่อยได้แก่ Training_Set, Validation_Set, Test_Set เพื่อจัดเก็บภาพสำหรับแต่ละประเภทเพื่อใช้ในการทดลอง

- ปรับขนาดภาพเป็น 224 X 224 พิกเซล (Pixel) ให้สอดคล้องกับแบบจำลองที่นำมาใช้ในการทดลอง

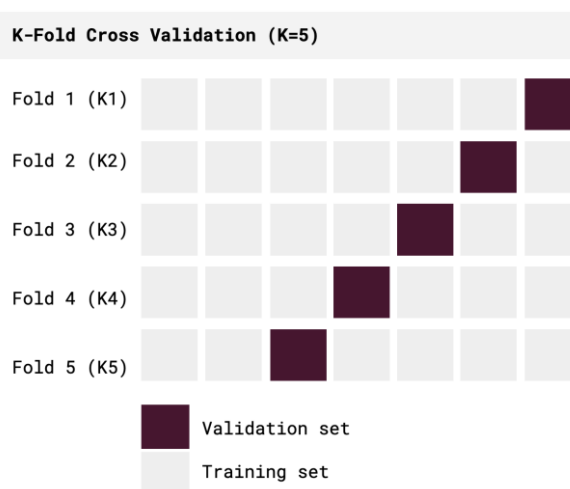
- ปรับค่ามาตรฐาน (Min-max normalization) โดยการสเกลค่าพิกเซลของภาพจากเดิมที่มีค่า 0 ถึง 255 ให้มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้ง่ายต่อการประมวลผลของโครงข่ายประสาทเทียม

- ส่วนป้ายกำกับคลาส (Labels) จะถูกแปลงให้อยู่ในรูปแบบ One-hot encoding เช่น [1, 0] สำหรับ Fractured และ [0, 1] สำหรับ Not fractured เพื่อใช้ในการคำนวณ Loss Function แบบ Categorical crossentropy

- การเรียนรู้ของแบบจำลองได้กำหนดพารามิเตอร์พื้นฐานในหัวข้อพารามิเตอร์พื้นฐานในการทดลอง (Hyper-parameters) ในทุกการทดลองและกำหนดรอบการเรียนรู้ 15 รอบ (Epoch)



(a)



(b)

ภาพที่ 2 (a) การแบ่งชุดข้อมูลสำหรับการเรียนรู้ของแบบจำลอง

(b) การแบ่งข้อมูล K-Fold Cross Validation (K=5)

การประเมินประสิทธิภาพแบบจำลอง

การวัดค่าความถูกต้อง (Accuracy) เป็นตัวชี้วัดที่ใช้สำหรับประเมินประสิทธิภาพของแบบจำลองจำแนกประเภท (Classification) โดยคิดเป็นสัดส่วนระหว่างจำนวนตัวอย่างที่แบบจำลองพยากรณ์ถูกต้องกับจำนวนตัวอย่างทั้งหมดในชุดข้อมูล (Chedsom, 2025) โดยใช้การคำนวณดังนี้

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

โดยที่

TP (True Positive) = คำตอบที่พยากรณ์ถูกและเป็นจริง

TN (True Negative) = คำตอบที่พยากรณ์ผิดแต่เป็นจริง

FP (False Positive) = พยากรณ์ผิดว่าเป็นจริง

FN (False Negative) = พยากรณ์ผิดว่าไม่เป็นจริง

การวัดค่าความแม่นยำ (Precision) ประเมินความถูกต้องและน่าเชื่อถือของการทำนายผล เป็นการวัดสัดส่วนของจำนวนตัวอย่างที่แบบจำลองทำนายได้ถูกต้อง (TP) ต่อจำนวนตัวอย่างที่แบบจำลองทำนายว่าเป็นจริงทั้งหมดทั้งที่ถูกและผิด (TP + FP) โดยใช้การคำนวณดังนี้

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

การวัดค่าความครบถ้วน (Recall) ประเมินความสามารถของแบบจำลองในการวัดสัดส่วนของจำนวนตัวอย่างที่แบบจำลองทำนายได้ถูกต้อง (TP) ต่อจำนวนตัวอย่างที่ทำนายว่าเป็นจริงทั้งหมดในชุดข้อมูลทั้งที่ทายถูกและทายผิดว่าปกติ (TP + FN) โดยใช้การคำนวณดังนี้

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

การวัดค่าความจำเพาะ (Specificity) ประเมินความสามารถของแบบจำลองในการระบุข้อมูลในคลาสที่ไม่ใช่เป้าหมายหลัก (Negative Class) ได้อย่างถูกต้อง คำนวณจากสัดส่วนของข้อมูลที่ทำนายว่าเป็นลบและถูกต้อง (TN) เปรียบเทียบกับจำนวนข้อมูลที่เป็นลบจริงทั้งหมดในชุดข้อมูล (TN + FP) โดยใช้การคำนวณดังนี้

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

การวัดค่าเอฟวันสกอร์ (F1-Score) ประเมินประสิทธิภาพรวมของแบบจำลองโดยพิจารณาความสมดุลระหว่างค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) เพื่อให้มีความสมดุลระหว่างการลดจำนวนผลบวกปลอม (FP) และผลลบปลอม (FN) โดยใช้การคำนวณดังนี้

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

การวัดความสูญเสีย (Loss Function) หรือฟังก์ชันความสูญเสีย ใช้วัด “ปริมาณความผิดพลาด” ของแบบจำลองในการพยากรณ์ผล ด้วยการคำนวณผลรวมความแตกต่างระหว่างค่าความน่าจะเป็นที่แบบจำลองทำนายได้กับผลเฉลยจริง (Actual Label) แบบจำลองที่มีประสิทธิภาพดีจะมีค่า Loss ลดลงและเข้าใกล้ค่าศูนย์ โดยใช้การคำนวณดังนี้

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (6)$$

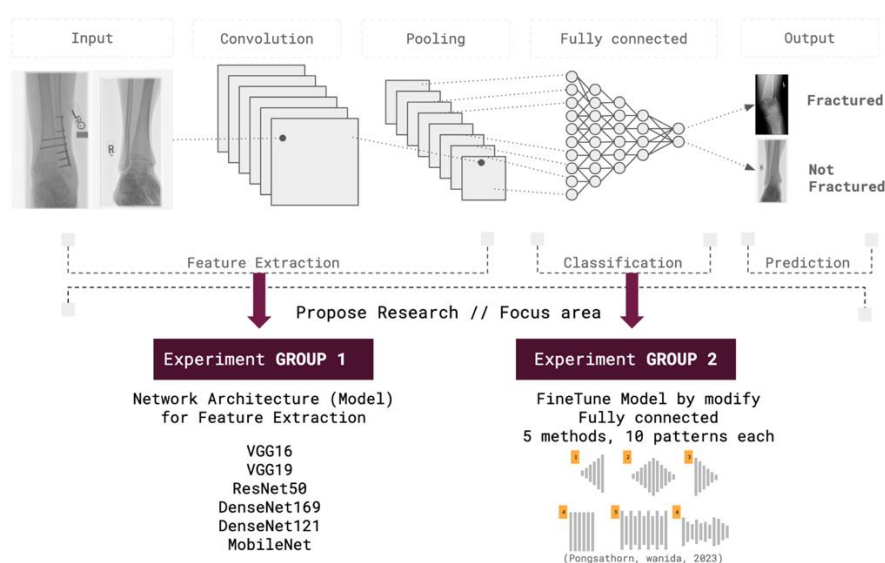
พารามิเตอร์พื้นฐานในการทดลอง (Hyper-Parameters)

ในการทดลองได้กำหนดพารามิเตอร์พื้นฐานสำหรับทุกๆ การทดลองเพื่อให้สภาวะแวดล้อมในการทดลองที่เหมือนกัน

- Batch size: 32
- Optimizer: Adam
- Learning rate: 0.001 (ค่าพื้นฐานของ Adam)
- Epoch: 15

การออกแบบการทดลอง

การทดลองแบ่งออกเป็น 2 กลุ่มโดยกลุ่มที่ 1 ใช้แบบจำลองพื้นฐานจำนวน 6 แบบจำลองเพื่อนำมาใช้ในการสกัดคุณลักษณะเด่นของข้อมูล (Feature extraction) ร่วมกับกลุ่มการทดลองที่ 2 ในการจำแนก (Classification) ข้อมูลในชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected layer) จากรูปแบบการปรับแต่งของ (Chedsom & Kanarkard, 2023) ดังการแสดงในภาพที่ 3 และแสดงขั้นตอนการทดลองกลุ่มที่ 1 และ กลุ่มที่ 2 ในภาพที่ 5 และแสดงภาพรวมการพัฒนาแบบจำลองในภาพที่ 6



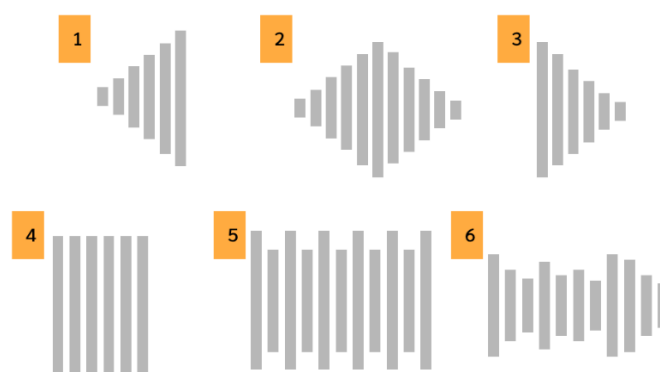
ภาพที่ 3 ภาพรวมการทดลอง

กลุ่มการทดลองที่ 1 แบบจำลองพื้นฐานจำนวน 6 แบบจำลอง

- VGG16 (Navale, S., 2023; Barua et al., 2023; Senapati et al., 2024)
- VGG19 (Barua et al., 2023; Senapati et al., 2024)
- ResNet50 (Senapati et al., 2024; Islam et al., 2025)
- DenseNet169 (Barua et al., 2023)
- DenseNet121 (Barua et al., 2023)
- MobileNet (Thota et al., 2024)

กลุ่มการทดลองที่ 2 ปรับแต่งชั้น Fully connected (ชั้นการเชื่อมโยงข้อมูลอย่างสมบูรณ์) จากรูปแบบการปรับแต่งของ (Chedsom & Kanarkard, 2023) 6 รูปแบบ รูปแบบละ จำนวน 10 ตัวอย่าง ดังแสดงในภาพที่ 4 ได้แก่

- รูปแบบที่ 1 การเพิ่มขึ้นของจำนวน NN (Neuron Network) สม่่าเสมอ ดังแสดงในตารางที่ 1
- รูปแบบที่ 2 การเพิ่มขึ้นของจำนวน NN (Neuron Network) และลดลงอย่างสม่่าเสมอเท่ากัน ดังแสดงในตารางที่ 1
- รูปแบบที่ 3 การลดจำนวนของ NN (Neuron Network) สม่่าเสมอ ดังแสดงในตารางที่ 2
- รูปแบบที่ 4 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) เท่าเดิม (คงที่) ดังแสดงในตารางที่ 2
- รูปแบบที่ 5 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) สลับกัน (สองค่า) ดังแสดงในตารางที่ 3
- รูปแบบที่ 6 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) เพิ่มและลดไม่เท่ากัน ดังแสดงในตารางที่ 3



ภาพที่ 4 วิธีการปรับแต่ง Neuron Network ใน Fully connected Layer 6 รูปแบบ (Chedsom & Kanarkard, 2023)

ตารางที่ 1 จำนวนโหนดของ Neuron Network ในรูปแบบที่ 1 – 2

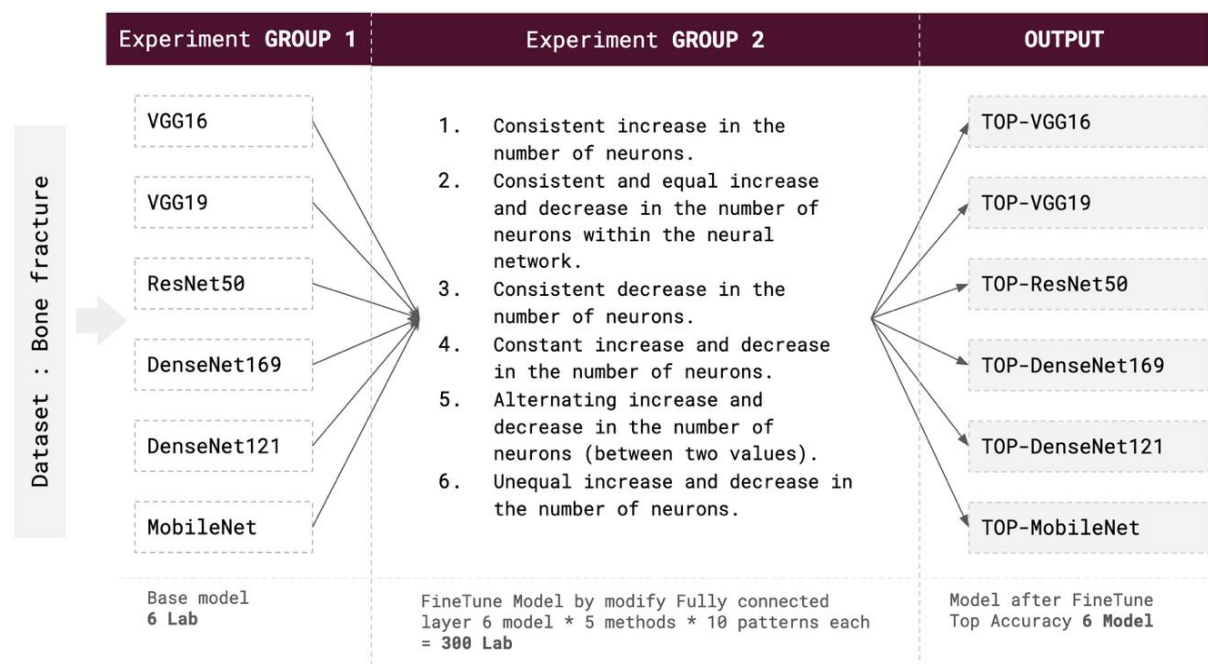
รูปแบบ	รูปแบบที่ 1 การเพิ่มขึ้นของจำนวน NN (Neuron Network) สม่่าเสมอ	รูปแบบ	รูปแบบที่ 2 การเพิ่มขึ้นของจำนวน NN (Neuron Network) และลดลงอย่าง สม่่าเสมอเท่ากัน
1-1	4, 8, 16, 32, 64, 128, 256, 512, 1024	2-1	4, 8, 16, 32, 64, 32, 16, 8, 4
1-2	8, 16, 32, 64, 128, 256, 512, 1024	2-2	8, 16, 32, 64, 128, 64, 32, 16, 8
1-3	16, 32, 64, 128, 256, 512, 1024	2-3	16, 32, 64, 128, 256, 128, 64, 32, 16
1-4	32, 64, 128, 256, 512, 1024	2-4	32, 64, 128, 256, 512, 256, 128, 64, 32
1-5	4, 16, 64, 256, 1024	2-5	4, 16, 64, 256, 1024, 256, 64, 16, 4
1-6	8, 32, 128, 512, 1024	2-6	8, 32, 128, 512, 1024, 512, 128, 32, 8
1-7	4, 8, 16, 32, 64, 128, 256, 512, 1024	2-7	4, 8, 16, 32, 64, 128, 64, 32, 16, 8
1-8	16, 64, 256, 1024	2-8	16, 64, 256, 1024, 256, 64, 16
1-9	4, 16, 64, 256, 1024	2-9	4, 16, 64, 256, 1024, 256, 64, 16, 4
1-10	8, 32, 128, 512, 1024	2-10	8, 32, 128, 512, 1024, 512, 128, 32, 8

ตารางที่ 2 จำนวนโหนดของ Neuron Network ในรูปแบบที่ 3 – 4

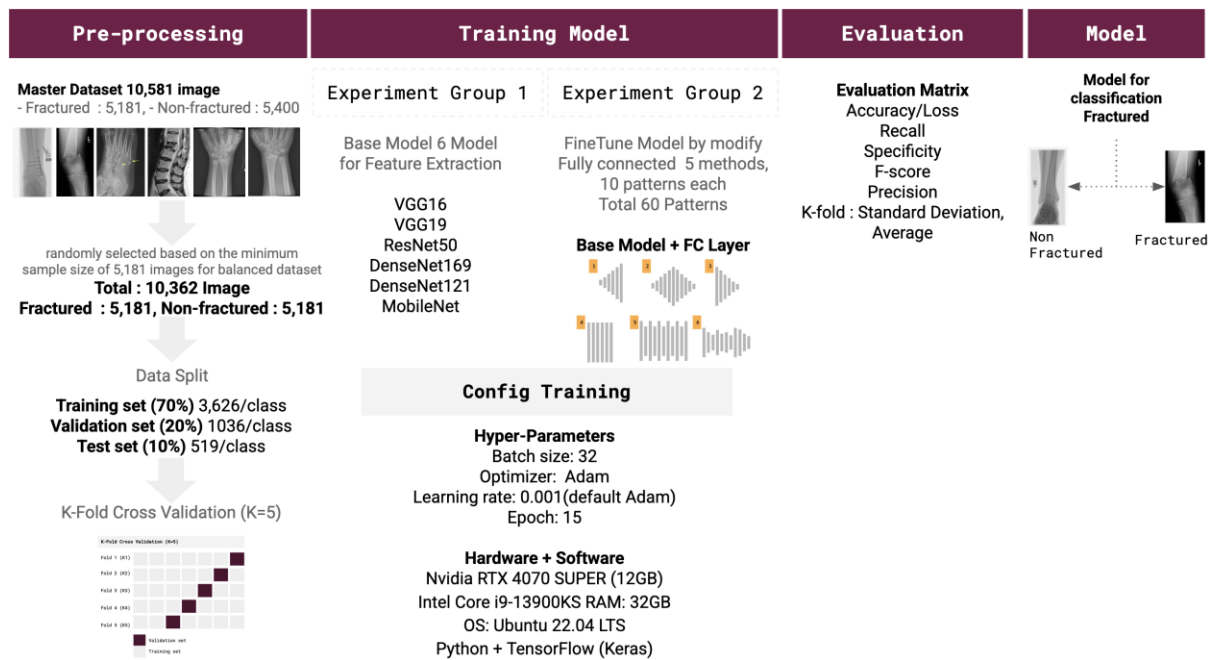
รูปแบบ	รูปแบบที่ 3 การลดจำนวนของ NN (Neuron Network) สม่่าเสมอ	รูปแบบ	รูปแบบที่ 4 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) เท่าเดิม (คงที่)
3-1	1024, 512, 256, 128, 64, 32, 16, 8, 4	4-1	64, 64, 64, 64, 64, 64, 64, 64, 64
3-2	1024, 512, 256, 128, 64, 32, 16, 8	4-2	128, 128, 128, 128, 128, 128, 128, 128, 128
3-3	512, 256, 128, 64, 32, 16, 8, 4	4-3	256, 256, 256, 256, 256, 256, 256, 256, 256
3-4	256, 128, 64, 32, 16, 8, 4	4-4	512, 512, 512, 512, 512, 512, 512, 512, 512
3-5	1024, 256, 64, 16, 4	4-5	1024, 1024, 1024, 1024, 1024, 1024, 1024, 1024, 1024
3-6	1024, 512, 128, 32, 8	4-6	4, 4, 4, 4, 4, 4, 4, 4, 4
3-7	1024, 512, 256, 128, 64, 32, 16, 8, 4	4-7	8, 8, 8, 8, 8, 8, 8, 8, 8
3-8	1024, 256, 64, 16	4-8	16, 16, 16, 16, 16, 16, 16, 16, 16
3-9	1024, 256, 64, 16, 4	4-9	32, 32, 32, 32, 32, 32, 32, 32, 32
3-10	1024, 512, 128, 32, 8	4-10	64, 64, 64, 64, 64, 64, 64, 64, 64

ตารางที่ 3 จำนวนโหนดของ Neuron Network ในรูปแบบที่ 5 – 6

รูปแบบ	รูปแบบที่ 5 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) สลับกัน (สองค่า)	รูปแบบ	รูปแบบที่ 6 การเพิ่ม-ลด จำนวนของ NN (Neuron Network) เพิ่มและลดไม่เท่ากัน
5-1	4, 8, 4, 8, 4, 8, 4, 8, 4, 8	6-1	4, 16, 64, 32, 128, 8, 256, 4, 512, 16
5-2	16, 32, 16, 32, 16, 32, 16, 32, 16, 32	6-2	8, 32, 128, 64, 256, 16, 512, 8, 1024, 32
5-3	64, 128, 64, 128, 64, 128, 64, 128, 64, 128	6-3	16, 64, 256, 128, 512, 32, 1024, 16, 4, 64
5-4	256, 512, 256, 512, 256, 512, 256, 512, 256, 512	6-4	32, 128, 512, 256, 1024, 64, 4, 32, 8, 128
5-5	1024, 4, 1024, 4, 1024, 4, 1024, 4, 1024, 4	6-5	64, 256, 1024, 512, 4, 128, 8, 64, 16, 256
5-6	8, 16, 8, 16, 8, 16, 8, 16, 8, 16	6-6	128, 512, 4, 256, 16, 32, 64, 1024, 8, 32
5-7	32, 64, 32, 64, 32, 64, 32, 64, 32, 64	6-7	256, 1024, 8, 128, 32, 64, 16, 512, 4, 128
5-8	128, 256, 128, 256, 128, 256, 128, 256, 128, 256	6-8	512, 4, 16, 64, 128, 32, 256, 8, 1024, 64
5-9	512, 1024, 512, 1024, 512, 1024, 512, 1024, 512, 1024	6-9	1024, 8, 32, 128, 256, 64, 512, 16, 4, 256
5-10	4, 1024, 4, 1024, 4, 1024, 4, 1024, 4, 1024	6-10	4, 32, 64, 256, 512, 128, 1024, 32, 8, 512



ภาพที่ 5 ภาพรวมขั้นตอนการทดลองกลุ่มที่ 1 -2



ภาพที่ 6 ภาพรวมขั้นตอนการพัฒนาแบบจำลอง

ผลการวิจัย

การทดลองแบ่งออกเป็น 2 กลุ่มการทดลอง

กลุ่มการทดลองที่ 1 แบบจำลองพื้นฐานจำนวน 6 แบบจำลอง พบว่า แบบจำลองที่ได้ผลดีที่สุดคือ VGG16 และ MobileNet ที่ให้ค่าความถูกต้อง (Accuracy) ร้อยละ 99.450 และมีค่าความสูญเสีย (Loss) 0.030 ทั้ง 6 แบบจำลองได้ผลค่าความถูกต้องไม่น้อยกว่าร้อยละ 90 ดังแสดงในตารางที่ 4

ตารางที่ 4 ผลการทดลองกลุ่มที่ 1 แบบจำลองพื้นฐานจำนวน 6 แบบจำลอง

No.	Network Architecture	Accuracy (Acc.)			Loss		
		Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	99.450	0.0004	99.69	0.030	0.002	0.020
2	VGG19	99.410	0.0005	99.29	0.014	0.003	0.010
3	ResNet50	93.280	0.0006	92.03	0.191	0.005	0.110
4	DenseNet169	98.620	0.0004	97.20	0.054	0.002	0.012
5	DenseNet121	99.210	0.001	98.40	0.031	0.004	0.040
6	MobileNet	99.450	0.0004	97.50	0.030	0.005	0.050

กลุ่มการทดลองที่ 2 ปรับแต่งชั้น Fully connected (ชั้นการเชื่อมโยงของข้อมูล) 6 รูปแบบ
 รูปแบบที่ 1 การเพิ่มขึ้นของจำนวน Neuron Network (NN) สมำเสมอ พบว่า แบบจำลองที่ได้ผลดีที่สุดคือแบบจำลอง VGG16 ร่วมกับ Fully connected Layer (FC) Patten 4 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.0000074) และ MobileNet ร่วมกับ Fully connected layer (FC) Patten 10 Fully connected layer (FC) Patten 3 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.000202) สูงที่สุดในรูปแบบที่ 1 ดังแสดงในตารางที่ 5

ตารางที่ 5 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 1 การเพิ่มขึ้นของจำนวน Neuron Network (NN) สมำเสมอ

No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	4	100.00	0.008	99.59	0.000 (0.0000074)	0.01	0.01
2	VGG19	8	99.60	0.01	90.05	0.021	0.06	0.09
3	ResNet50	10	52.96	0.04	91.65	0.693	0.07	0.29
4	DenseNet169	8	99.41	0.12	88.86	0.026	0.21	0.24
5	DenseNet121	7	99.21	0.08	93.78	0.046	0.14	0.13
6	MobileNet	10	100.00	0.04	97.66	0.000 (0.000202)	0.04	0.04

รูปแบบที่ 2 การเพิ่มขึ้นของจำนวน Neuron Network (NN) และลดลงอย่างสมำเสมอเท่ากัน พบว่า แบบจำลอง VGG16 ร่วมกับ Fully connected layer (FC) Patten 8 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.0000200) และ MobileNet ร่วมกับ Fully connected layer (FC) Patten 10 Fully connected layer (FC) Patten 3 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.002 สูงที่สุดในรูปแบบที่ 2 ดังแสดงในตารางที่ 6

ตารางที่ 6 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 2 การเพิ่มขึ้นของจำนวน Neuron Network (NN) และลดลงอย่างสม่ำเสมอเท่ากัน

No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	3	100.00	0.004	99.92	0.000 (0.0000200)	0.01	0.01
2	VGG19	1	99.60	0.08	93.87	0.029	0.21	0.16
3	ResNet50	3	92.88	0.19	84.71	0.211	0.25	0.23
4	DenseNet169	9	99.40	0.04	92.07	0.036	0.17	0.20
5	DenseNet121	6	99.20	0.04	96.12	0.078	0.11	0.12
6	MobileNet	10	100.00	0.00	99.45	0.002	0.03	0.03

รูปแบบที่ 3 การลดจำนวนของ Neuron Network (NN) สม่ำเสมอ พบว่า แบบจำลอง VGG16 ร่วมกับ Fully connected Layer (FC) Patten 7 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.00000131) สูงที่สุดในรูปแบบที่ 3 ดังแสดงในตารางที่ 7

ตารางที่ 7 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 3 การลดจำนวนของ Neuron Network (NN) สม่ำเสมอ

No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	7	100.00	0.22	90.03	0.000 (0.00000131)	0.30	0.14
2	VGG19	2	99.41	0.06	95.32	0.006	0.21	0.11
3	ResNet50	2	93.67	0.01	96.93	0.169	0.03	0.10
4	DenseNet169	4	99.20	0.27	80.07	0.068	0.37	0.27
5	DenseNet121	8	99.20	0.02	99.97	0.038	0.00	0.00
6	MobileNet	8	99.40	0.02	99.97	0.018	0.00	0.00

รูปแบบที่ 4 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) เท่าเดิม (คงที่) พบว่า แบบจำลอง VGG16 ร่วมกับ Fully connected Layer (FC) Patten 3 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.000000172) ดังแสดงโครงสร้างในภาพที่ 8 ที่ให้ประสิทธิภาพมากที่สุดและแบบจำลอง MobileNet ร่วมกับ Fully connected Layer (FC) Patten 1 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.008 สูงที่สุดในรูปแบบที่ 4 ดังแสดงในตารางที่ 8

ตารางที่ 8 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 4 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) เท่าเดิม (คงที่)

No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	3	100.00	0.01	99.95	0.000 (0.000000172)	0.04	0.04
2	VGG19	1	99.40	0.01	99.92	0.005	0.04	0.04
3	ResNet50	1	93.47	0.01	96.37	0.172	0.04	0.04
4	DenseNet169	1	99.40	0.00	99.95	0.030	0.00	0.02
5	DenseNet121	2	99.20	0.04	99.97	0.028	0.11	0.07
6	MobileNet	1	100.00	0.04	99.97	0.008	0.07	0.07

รูปแบบที่ 5 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) สลับกัน (สองค่า) พบว่าแบบจำลอง VGG16 ร่วมกับ Fully connected Layer (FC) Patten 8 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.00000216) และแบบจำลอง DenseNet169 ร่วมกับ Fully connected Layer (FC) Patten 6 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.000629) และแบบจำลอง MobileNet ร่วมกับ Fully connected Layer (FC) Patten 4 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.0000111) สูงที่สุดในรูปแบบที่ 5 ดังแสดงในตารางที่ 9

ตารางที่ 9 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 5 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) สลับกัน (สองค่า)

No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	8	100.00	0.04	97.94	0.000 (0.00000216)	0.11	0.08
2	VGG19	2	99.40	0.04	97.95	0.029	0.12	0.07
3	ResNet50	2	93.87	0.08	94.02	0.169	0.14	0.13
4	DenseNet169	6	100.00	0.08	95.83	0.000 (0.000629)	0.12	0.08
5	DenseNet121	2	99.20	0.08	93.69	0.044	0.20	0.17
6	MobileNet	4	100.00	0.08	93.56	0.000 (0.0000111)	0.17	0.17

รูปแบบที่ 6 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) เพิ่มและลดไม่เท่ากัน พบว่าแบบจำลอง VGG16 ร่วมกับ Fully connected layer (FC) Patten 5 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.00 (0.00000217) และแบบจำลอง DenseNet169 ร่วมกับ Fully connected layer (FC) Patten 4 ให้ค่าความถูกต้องร้อยละ 100 และให้ค่าความสูญเสีย 0.000 (0.00000434) สูงที่สุดในรูปแบบที่ 6 ดังแสดงในตารางที่ 10

ตารางที่ 10 ผลการทดลองกลุ่มที่ 2 รูปแบบที่ 6 การเพิ่ม-ลด จำนวนของ Neuron Network (NN) เพิ่มและลดไม่เท่ากัน 17 18

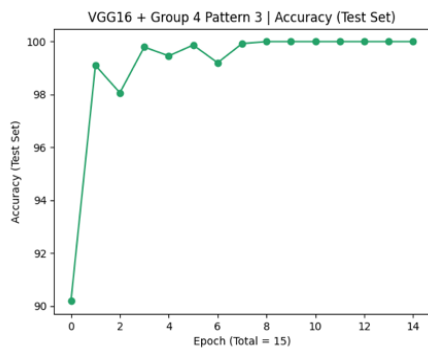
No.	Network Architecture	Patten	Accuracy (Acc.)			Loss		
			Acc.	SD.	Average	Loss	SD.	Average
1	VGG16	5	100.00	0.07	98.63	0.000	0.18	0.21
						(0.00000217)		
2	VGG19	2	99.40	0.05	95.55	0.027	0.19	0.16
3	ResNet50	10	52.96	0.20	65.70	0.692	0.40	0.45
4	DenseNet169	3	98.61	0.21	87.88	0.100	0.28	0.19
5	DenseNet121	4	100.00	0.21	88.04	0.000	0.29	0.18
						(0.00000434)		
6	MobileNet	8	100.00	0.22	89.98	0.001	0.15	0.15

การเรียนรู้ของแบบจำลอง VGG16 ร่วมกับ Fully connected layer (FC) กลุ่มที่ 4 รูปแบบที่ 3 แสดงในภาพที่ 7 (a) กราฟค่าความถูกต้องของการเรียนรู้กับชุดข้อมูล Test set ได้แสดงให้เห็นถึงความถูกต้องสูงสุดตั้งแต่รอบการเรียนรู้ที่ 7 เป็นต้นไปอย่างคงที่ และภาพที่ 7 (b) เป็นกราฟค่าความสูญเสียของการเรียนรู้ที่ลดลงอย่างสม่ำเสมอตั้งแต่รอบที่ 1 และคงที่ในรอบที่ 8 เป็นต้นไป อีกทั้งการแสดงตาราง Confusion matrix ในภาพที่ 7 (c) ได้แสดงจำนวนการจำแนกภาวะกระดูกหัก (Fractured) ถูกต้องทั้งหมดจาก 519 ภาพ และสภาวะปกติกระดูกไม่หัก (Non-Fractured) ถูกต้องทั้งหมด 519 ภาพ และในการวัดค่าประสิทธิภาพได้แก่การวัดค่าความแม่นยำ (Precision) มีค่าร้อยละ 100.00 และการวัดค่าความจำเพาะ (Specificity) มีค่าร้อยละ 100.00 และในการวัดค่าความครบถ้วน (Recall / Sensitivity) มีค่าร้อยละ 100.00 และมีค่า F1-Score ร้อยละ 100.00

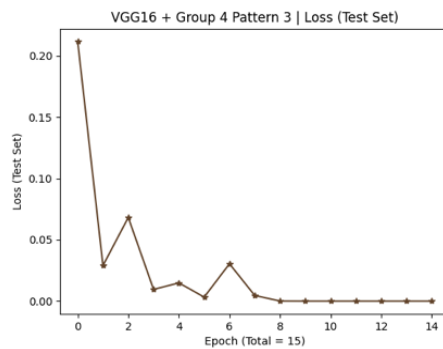
จากตารางเปรียบเทียบจะเห็นได้ว่างานวิจัยนี้มีระดับความแม่นยำที่สูงที่สุดซึ่งเกิดจากการปรับแต่งโครงสร้างชั้นเชื่อมโยงแบบสมบูรณ์ (Fully connected layer) อย่างเป็นระบบถึง 60 รูปแบบ ขณะที่งานวิจัยของ (Navale, 2023) หรือ (Barua et al., 2023) มักใช้แบบจำลองมาตรฐานโดยไม่มีการปรับแต่งโครงสร้างส่วนปลายอย่างละเอียดทำให้ได้ค่าความถูกต้องในช่วงร้อยละ 79 - 88 อีกทั้งงานวิจัยอื่นอาจใช้ชุดข้อมูลที่มีความท้าทายหรือมีความซับซ้อนของรอยโรคที่ต่างออกไป ในงานวิจัยนี้ให้ค่าความถูกต้องร้อยละ 100 สะท้อนถึงความเข้ากันได้ (Optimal fit) ระหว่างแบบจำลองที่พัฒนาขึ้นกับชุดข้อมูลกระดูกหักที่เลือกมาใช้อย่างเหมาะสม

ตารางที่ 11 การเปรียบเทียบประสิทธิภาพการจำแนกภาวะกระดูกหักกับงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้อง	แบบจำลอง	ชุดข้อมูล	ค่าความถูกต้อง
(Navale, 2023)	VGG16	MURA-v1.1	85.82%
(Barua et al., 2023)	Ensemble (VGG16+VGG19)	MURA	83.00%
(Senapati et al., 2024)	VGG16	ภาพเอกซเรย์ข้อมือ	84.00%
(Islam et al., 2025)	Visual Transformers (ViT)	MURA	88.00%
(Thota et al., 2024)	MobileNet	ภาพเอกซเรย์ทั่วไป	98.00%
งานวิจัยนี้	VGG16 + Optimized FC Layer	Bone Fracture Multi-Region	100%



(a) กราฟค่าความถูกต้องของการเรียนรู้

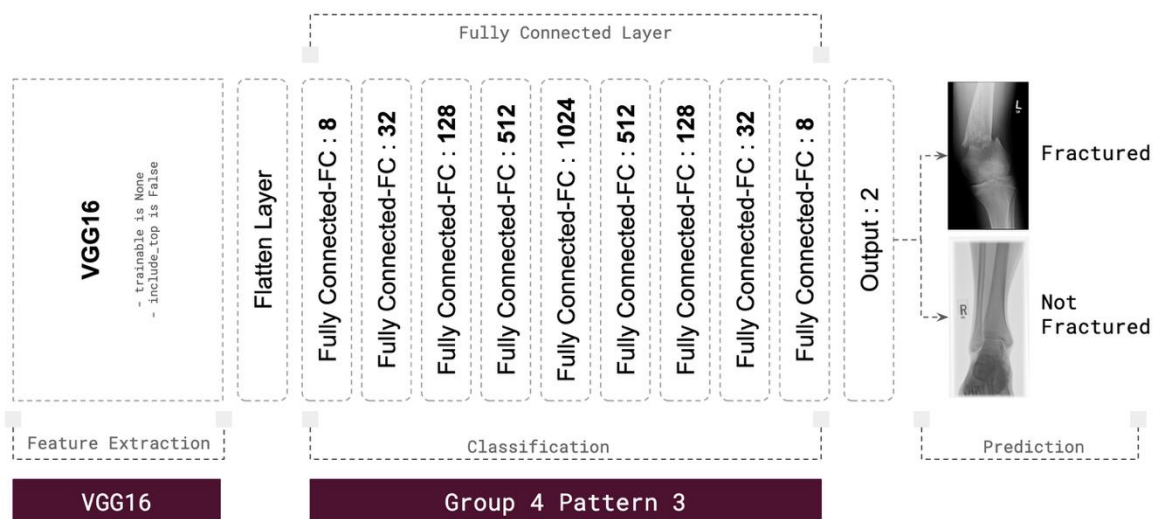


(b) กราฟค่าความสูญเสียของการเรียนรู้

	Fractured	Non Fractured		
	Fractured	Non Fractured	Recall	Precision
			100	100
Non Fractured	0	519	Specificity	F1-score
	0	519	100	100

(c) ตาราง Confusion Matrix และการวัดประสิทธิภาพ

ภาพที่ 7 ผลการเรียนรู้ของแบบจำลอง



ภาพที่ 8 โครงสร้างแบบจำลองที่ให้ผลดีที่สุด (VGG16 + กลุ่มที่ 4 รูปแบบที่ 3)

การอภิปรายผล

แบบจำลองที่ให้ผลดีที่สุดในการทดลองคือแบบจำลอง VGG16 (Simonyan & Zisserman, 2016) ร่วมกับ Fully connected Layer (FC) กลุ่มที่ 4 รูปแบบที่ 3 (Neurons : 256, 256, 256, 256, 256, 256, 256, 256) ให้ค่าความถูกต้อง (Accuracy) ร้อยละ 100 ค่าความสูญเสีย (Loss) ร้อยละ 0.000000172 (ค่าความสูญเสียน้อยที่สุด) แสดงโครงสร้างแบบจำลองในภาพที่ 8 การเรียนรู้ของแบบจำลองนี้เพื่อการป้องกันการ Overfitting ในงานวิจัยนี้มีการแบ่งข้อมูลในปริมาณที่เท่ากันในลักษณะ Balanced dataset โดยมีภาพภาวะกระดูกหักจำนวน 5,216 ภาพ และภาพปกติจำนวน 5,216 ภาพ ในสัดส่วนร้อยละ 50:50 ช่วยลดโอกาสที่แบบจำลองจะมีอคติ (Bias) ต่อกลุ่มข้อมูลด้านใดด้านหนึ่งซึ่งเป็นปัจจัยสำคัญที่ส่งผลให้ค่าความถูกต้อง (Accuracy) สะท้อนความถูกต้องของแบบจำลองที่แท้จริง และใช้เทคนิค K-Fold cross validation K=5 เพื่อให้แบบจำลองมีผลลัพธ์ที่เสถียรและสม่ำเสมอในทุกส่วนของข้อมูล ซึ่งได้รายงานค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน (SD) นี้ด้วย และการวิเคราะห์ค่าความสูญเสียจากกราฟผลการทดลองในภาพที่ 7 (b) พบว่าค่าความสูญเสียลดลงอย่างต่อเนื่องจนเข้าใกล้ศูนย์ (0.000000172) และคงที่ โดยไม่มีการติดกลับของค่า Loss ในชุด Validation/Test ซึ่งแสดงให้เห็นว่าแบบจำลองมีการเรียนรู้ที่สมบูรณ์และไม่มีภาวะ Overfitting เกิดขึ้น

จากการทดลองในกลุ่มที่ 1 แบบจำลองที่ให้ผลดีที่สุดคือ VGG16 และ MobileNet ค่าความถูกต้องร้อยละ 100 กลุ่มการทดลองที่ 2 มีแบบจำลองที่ให้ผลดีที่สุดให้ค่าความถูกต้องร้อยละ 100 จำนวน 22 แบบจำลองแบ่งเป็น VGG16 จำนวน 12 แบบจำลอง MobileNet จำนวน 8 แบบจำลอง ส่วน DenseNet169 และ DenseNet121 อย่างละ 1 แบบจำลอง แสดงให้เห็นว่าแบบจำลอง VGG16 ร่วมกับการปรับแต่งชั้น Fully connected Layer (FC) แต่ละรูปแบบให้ความสามารถในการจำแนกภาวะกระดูกหักจากภาพเอกซเรย์กระดูกข้อมือได้ผลดีจำนวนมากที่สุด อีกนัยยะคือแบบจำลองนี้สามารถเข้ากับชุดข้อมูลนี้มากที่สุด

โครงสร้างของแบบจำลองที่นำมาใช้ในการสกัดคุณลักษณะเด่นของข้อมูลทั้งหมด 6 แบบจำลอง VGG16 มีความโดดเด่นคือการใช้ตัวกรองขนาดเล็ก (3×3) ซ้อนกันหลายชั้นทำให้สกัดคุณลักษณะเด่นที่มีความละเอียดสูง (Fine-grained features) ได้อย่างมีประสิทธิภาพ ซึ่งสอดคล้องกับลักษณะเฉพาะของรอยแตกหักในภาพเอกซเรย์กระดูกรยางค์ที่มีลักษณะเป็นเส้นริ้วขนาดเล็กหรือความผิดปกติของพื้นผิวขอบกระดูกที่มีความซับซ้อนการใช้ตัวกรองขนาดเล็กซ้อนกันช่วยให้แบบจำลองตรวจจับรายละเอียดความต่อเนื่องของรอยโรคได้ดีกว่าตัวกรองที่มีขนาดใหญ่ MobileNet มีข้อได้เปรียบเรื่องความเร็วและขนาดเล็ก (Lightweight) ทำให้เหมาะสำหรับการพัฒนาเป็นแอปพลิเคชันบนอุปกรณ์พกพาหรือ Edge Computing โดยให้ค่าความถูกต้องสูงถึงร้อยละ 100 เมื่อปรับแต่ง Fully connected layer (FC) ส่วน ResNet50 และ DenseNet เป็นสถาปัตยกรรมที่ทันสมัยมากกว่าด้วยเทคนิคการข้ามชั้น (Residual Connection) หรือการรวมคุณลักษณะซ้ำ (Feature reuse) แต่ในการทดลองนี้ค่าความถูกต้องที่น้อยกว่า โดยเฉพาะ ResNet50 ที่มีค่าความถูกต้องต่ำสุดในบางการทดลอง ซึ่งอาจเกิดจากความซับซ้อนของแบบจำลองที่มีมากเกิดความจำเป็น (Over-complexity) สำหรับลักษณะข้อมูลภาพนี้

เมื่อนำมาเปรียบเทียบกับผลการทดลองในบทความอื่น ๆ เช่นแบบจำลอง VGG16 ของ (Navale, 2023) ให้ค่าความถูกต้องร้อยละ 85.82 แบบจำลอง Ensemble-1 (VGG16 และ VGG19) ของ (Barua et al., 2023) มีความถูกต้องร้อยละ 83.00 และแบบจำลอง VGG16 มีความถูกต้องร้อยละ 79.00 แบบจำลอง VGG16 ของ (Senapati et al., 2024) มีความถูกต้องร้อยละ 84.00 และเปรียบเทียบแบบจำลองที่ได้ผลดีที่สุด (VGG16 + Group 4 Pattern 3) พบว่าค่าความถูกต้องเพิ่มขึ้นร้อยละ 0.55 และค่าความสูญเสียลดลงร้อยละ 0.0299 จากกลุ่มการทดลองที่ 1 แบบจำลองพื้นฐานและกลุ่มการทดลองที่ 2 ที่ปรับแต่งแบบจำลอง

งานวิจัยส่วนใหญ่ เช่น (Barua et al., 2023) หรือ (Navale, 2023) ที่เน้นการใช้แบบจำลองมาตรฐาน (Baseline) หรือการทำ Ensemble ซึ่งในงานวิจัยนี้ได้ทดลองหาโครงสร้างชั้นเชื่อมโยงแบบสมบูรณ์ 60 รูปแบบ เพื่อหาองค์ประกอบของ Fully connected Layer (FC) ที่เหมาะสมที่สุด (Optimal architecture) และผลการทดลองที่ให้ค่าความถูกต้องร้อยละ 100 สูงกว่างานของ (Navale, 2023) ที่ได้ร้อยละ 85.82 หรือ (Islam et al., 2025) ที่ใช้ Visual transformers แล้วได้ร้อยละ 88.00 แสดงให้เห็นว่าการปรับแต่ง FC layer มีผลสำคัญต่อความแม่นยำมากกว่าการเปลี่ยนตัวสกัดคุณลักษณะพื้นฐานเพียงอย่างเดียว

บทสรุปและข้อเสนอแนะ

จากงานวิจัยนี้นำเสนอกระบวนการพัฒนาแบบจำลองเพื่อทดสอบประสิทธิภาพของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) สำหรับการจำแนกภาวะกระดูกหักจากภาพเอกซเรย์กระดูกรยางค์ โดยแบ่งการทดลองออกเป็น 2 กลุ่มการทดลองกลุ่มที่ 1 ได้ทดลองโดยใช้แบบจำลองพื้นฐาน 6 แบบจำลองทำหน้าที่ในการสกัดคุณลักษณะเด่นจากภาพ (Feature extraction) และนำไปใช้การทดลองกลุ่มที่ 2 ในการจำแนกข้อมูล

(Classification) ที่แบ่งบอกถึงสถานะกระดูกหักหรือไม่หักในชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-connected layer) ที่มีการปรับเปลี่ยนโครงสร้างใน 6 กลุ่ม กลุ่มละ 10 รูปแบบ

ผลการทดลองจากงานวิจัยนี้แสดงให้เห็นว่าแบบจำลอง VGG16 + Group 4 Pattern 3 (Fully connected Layer (Neurons) : 256, 256, 256, 256, 256, 256, 256, 256, 256) จากกลุ่มการทดลองที่ 2 ให้ค่าความถูกต้อง (Accuracy) ร้อยละ 100 และค่าความสูญเสีย (Loss) ที่ต่ำมากที่สุดที่ 0.000000172 สามารถจำแนกภาวะกระดูกหักจากภาพเอกซเรย์กระดูกรยางค์ได้อย่างมีประสิทธิภาพกับชุดข้อมูลที่นำมาใช้ในการเรียนรู้และทดลองนี้

ในการทดลองเพื่อการพัฒนาแบบจำลองในอนาคตสามารถนำวิธีการนี้ไปใช้พัฒนาเครื่องมือช่วยวินิจฉัยภาวะกระดูกหักแบบอัตโนมัติจากภาพเอกซเรย์กระดูกรยางค์ได้ โดยเน้นที่การปรับแต่งโครงสร้างชั้นเชื่อมโยงแบบสมบูรณ์ (Fully connected layer) รูปแบบการจัดวาง Neuron network ทั้ง 60 รูปแบบที่นำเสนอในงานวิจัยนี้นำไปใช้ออกแบบสถาปัตยกรรมโครงข่ายประสาทเทียมเพื่อเพิ่มประสิทธิภาพการจำแนกภาพทางการแพทย์ที่มีความซับซ้อนสูงเพื่อให้แบบจำลองสอดคล้องกับชุดข้อมูลมากที่สุด รวมทั้งสนับสนุนการทำงานของรังสีแพทย์หรือลดภาระงานวินิจฉัยลงในการวินิจฉัยเบื้องต้น (Preliminary diagnosis) โดยการพัฒนาเครื่องมือสนับสนุนการวินิจฉัยทางคลินิกจากแบบจำลองที่พัฒนาขึ้น และนำไปประยุกต์ใช้กับชุดข้อมูลทางการแพทย์อื่น ๆ เช่น ภาพเอกซเรย์ของอวัยวะหรือบริเวณอื่น ๆ หรือประเภทโรคภาวะอื่น ๆ อีกทั้งเพื่อให้รองรับการใช้งานในทรัพยากรจำกัดด้วยแบบจำลองที่มีขนาดเล็ก (Lightweight) เช่น MobileNet มีความแม่นยำและใช้ทรัพยากรในการคำนวณที่ต่ำสามารถนำไปพัฒนาเป็นแอปพลิเคชันบนอุปกรณ์พกพาหรืออุปกรณ์ Edge computing ได้ นอกจากนี้เพื่อให้แบบจำลองได้เรียนรู้ข้อมูลที่มีความหลากหลายหลายควรใช้เทคนิคการเพิ่มปริมาณของข้อมูล (Data augmentation) และข้อมูลจากชุดข้อมูลอื่น ๆ มาใช้ในการเรียนรู้และการทดลองแบบจำลองเพื่อให้มีประสิทธิภาพมากยิ่งขึ้น

ข้อจำกัด (Limitation) ในการใช้งานและการทดลองในงานวิจัยนี้เนื่องจากแบบจำลองที่พัฒนาขึ้นนี้สามารถทำงานได้ตามผลการทดลองในสภาพแวดล้อมของชุดข้อมูลที่นำมาใช้ในการเรียนรู้และการทดลองเท่านั้นซึ่งหากการนำไปใช้ในสภาพแวดล้อมจริงควรมีการเพิ่มชุดข้อมูลที่มีความหลากหลายและมีปริมาณที่มากเพียงพอเพื่อให้เกิดความมั่นใจมากยิ่งขึ้นในทางการแพทย์ อีกทั้งชุดข้อมูลที่นำมาใช้ในการทดลองนี้ถูกนำมาจากฐานข้อมูลสาธารณะ (Kaggle) ซึ่งไม่ใช่ชุดข้อมูลที่เก็บรวบรวมโดยตรงจากโรงพยาบาลในประเทศไทย

เอกสารอ้างอิง

- Barua, K., Mahmud, T., Barua, A., Sharmen, N., Basnin, N., Islam, D., & Hossain, S. (2023). Explainable AI-based humerus fracture detection and classification from X-ray images. *In 2023 26th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, (pp. 1-6). IEEE.
<https://doi.org/10.1109/ICCIT60459.2023.10441124>
- Chedsom, P., & Kanarkard, W. (2023). Analysis of student engagement in online classroom using Convolutional Neural Networks (CNN). *ECTI Transaction on Application Research and Development*, 3(3), 39–52. <https://doi.org/10.37936/ectiard.2023-3-3.250499>
- Chedsom, P. (2025). Pneumonia Detection from Chest X-ray Images using Convolutional Neural Networks and Transfer Learning Techniques. *Journal of Applied Informatics and Technology*, 7(2), 406–431. <https://ph01.tci-thaijo.org/index.php/jait/article/view/255187>
- Islam, M. U., Fathima, A., & Ghosh, D. (2025). An improved technique for bone fracture classification using visual transformers and convolutional neural networks. *Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024)*. <http://dx.doi.org/10.2139/ssrn.5191555>
- Navale, S. (2023). Bone abnormalities detection and classification using deep learning-Vgg16 algorithm. *International Journal for Science Technology and Engineering*, 11(7), 122–129. <https://doi.org/10.22214/ijraset.2023.54582>
- Rodrigo, M. (2025). *Bone fracture multi-region X-ray data*.
<https://www.kaggle.com/datasets/bmadushanirodrigo/fracture-multi-region-x-ray-data>
- Sahin, M. E. (2023). Image processing and machine learning-based bone fracture detection and classification using X-ray images. *International Journal of Imaging Systems and Technology*, 33(3), 853–865. <https://doi.org/10.1002/ima.22849>
- Senapati, B., Naeem, A. B., Ghafoor, M. I., Gulaxi, V., Almeida, F., Anand, M. R., & Jaiswal, C. (2024). Wrist Crack Classification Using Deep Learning and X-Ray Imaging. In: Daimi, K., Al Sadoon, A. (eds). *Proceedings of the Second International Conference on Advances in Computing Research (ACR'24)*. ACR 2024. Lecture Notes in Networks and Systems, vol 956. Springer, Cham. https://doi.org/10.1007/978-3-031-56950-0_6

- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015, San Diego, 7-9 May 2015*, 1-14.
<https://arxiv.org/pdf/1409.1556.pdf>
- Thota, S., Kandukuru, P., Sundaram, M., Ali, A., Basha, S. M., & Bindu, N. H. (2024). "Deep Learning based Bone Fracture Detection," *2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, Tumakuru, India, (pp. 1-7).
<https://doi.org/10.1109/ICSSES62373.2024.10561360>